

# Social Networks through the Prism of Cognition

Radosław Michalski<sup>1\*</sup>, Bolesław K. Szymański<sup>2†</sup>, Przemysław Kazienko<sup>1</sup>,  
Christian Lebiere<sup>3</sup>, Omar Lizardo<sup>4</sup>, Marcin Kulisiewicz<sup>1</sup>

<sup>1</sup> Faculty of Computer Science and Management, Wrocław University of Science and Technology, Poland

<sup>2</sup> Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA

<sup>3</sup> Carnegie Mellon University, Pittsburgh, PA, USA

<sup>4</sup> University of Notre Dame, Notre Dame, IN, USA

## Abstract

Human relations are driven by social events - people interact, exchange information, share knowledge and emotions, or gather news from mass media. These events leave traces in human memory. The initial strength of a trace depends on cognitive factors such as emotions or attention span. Each trace continuously weakens over time unless another related event activity strengthens it. Here, we introduce a novel Cognition-driven Social Network (CogSNet) model that accounts for cognitive aspects of social perception and explicitly represents human memory dynamics. For validation, we apply our model to NetSense data on social interactions among university students. The results show that CogSNet significantly improves quality of modeling of human interactions in social networks.

**Keywords:** social network, network dynamics, memory model, cognition

## Discussion

There is a fundamental difference between the ways social events are currently represented by network analysts and the way they are perceived and cognitively processed by humans. In the former, the discrete nature of events is retained and the weights of social network edges are updated once per each relevant event. In human memory, by contrast, perception of events changes continuously over time. Moreover, the initial strength of a trace depends on cognitive factors defined by states of mind of participants, and specific aspects of their interactions. Decisions about whether to initiate, maintain or discontinue social relations involve cognitive processes operating on all relevant information stored in memory traces over specific time-scales.

To date, however, there are no models that accurately represent the dynamics of event-memory driven social relations. To address this gap, we introduce a novel Cognition-driven Social Network (CogSNet) model that captures impact of human memory on perception of accumulated events and on decisions to form, maintain, or dissolve social relations. The model explicitly represents human memory dynamics, such as the gradual decay of memory traces over time. With suitable data, it can be extended to include additional cognitive aspects, such as individual levels of sensitivity to relevant events, emotions or distractions during

---

\*E-mail:radoslaw.michalski@pwr.edu.pl

†E-mail:szymab@rpi.edu

perception of events. Hence, the model is capable of capturing the dynamics of social interactions in natural settings from the cognitive perspective of each participant.

To evaluate the empirical performance of the proposed model, we compare its predictions, based on behavioral communication event data, to the ground truth perceptual data collected from human participants with regard to their most important relations. The results reveal that the perception of the depth of interactions between people is well captured by the CogSNet model. At any given point in time, the model can compute the current strength of memory traces, including the impact of discrete events creating or reinforcing these traces. The computed state of each social relation (e.g. salient versus decayed) can then be compared against the externalization of cognition in the form of self-reports. The predictions of the CogSNet model achieve high accuracy for the NetSense network. This success demonstrates the importance of incorporating cognitive processes and memory dynamics for adequate modeling of the dynamics of social relationships.

Memory is considered to be one of the most important components of human cognition. This is especially the case given the necessity to efficiently retrieve large amounts of knowledge and to select information from a noisy environment. Hence, one of the fundamental challenges for cognitive science has been to understand the mechanisms involved in managing information in human memory [2]. The exact details of these mechanisms are difficult to firmly establish since the human brain is a highly complex system featuring strong differences across individuals based on experimental tuning [9]. Yet, there have been significant advances in the direction of developing good quality working models of memory and other core cognitive processes. For example, the ACT-R cognitive architecture's does a good job of modeling core features of human declarative memory, successfully replicating a large number of well-known effects. For our purposes, the most important among these empirical regularities are the well-established primacy and recency effects for list memory [1].

Limited capacity and the gradual decay of traces in memory over time have been confirmed by many studies analyzing the relationship between time and recall [10,17]. The forgetting mechanism is modeled by a parametric function describing how well a given item will be recalled as a function of time. It appears that humans tend to underestimate the number of events they experience, i.e., they actually forget faster than they think they do [10]. Obviously, any reminiscence of a particular event primed by an external situation allows people to remember the information longer. Yet ultimately, the limited capacity paradigm and the forgetting function are chiefly responsible for controlling the life time of such reminiscence in memory [13].

The dataset on which we operate contains data limited to a relatively small fraction of human social activities. Hence, it cannot be stated that there were no other events happening that could have affected participant memory about others. Examples of such events are face to face interactions, indirect references to someone else, and personal reminiscences about others. In addition, emotions can also play a significant role in these processes and they can be potentially identified with some confidence from available data [4].

As a result, the reinforcement peak value can be adjusted for a given person and for an individual event, cf. also Supplementary Eq. 6. What is observed here, is a partial manifestation from human memory. However, even taking into account that the model was built based on a single data source, over 6 million telephone calls and messages among the NetSense study participants, the model provides good accuracy in predicting the salience of social contacts over all 578 surveys completed by 184 participants. This accuracy most likely could have been increased if the parameters had been individually adjusted for each participant.

In future work, we plan to extend the model in the directions outlined in *Supplementary Information* including accounting for distractions during interactions, individualized strength, asymmetric of interactions of significance to participants (e.g., hierarchical relationships), and the impact of forms of interactions and of associated emotions. Hence, the CogSNet model represents an important first step towards modeling social network dynamics through the prism of human cognition.

## Methods

The human brain records events as they arrive, but only a small fraction of the incoming information is stored in long-term memory due to its limited capacity. The forgetting mechanism dictates that the chance to recall a given event decreases gradually as we move forward from the time of first exposure [17]. In some sense, it is similar to graph-streaming [6] and feed-based social media network cascades [15] scenarios where incoming events are ordered by their arrival time but only some of them are kept.

Accordingly, the CogSNet model uses a *forgetting function*  $f$  to account for the decreasing probability of keeping aging of memory traces over time. Forgetting is thus a monotonically non-increasing function of time with  $f(0) = 1$  and  $f(t) \geq 0$  for all  $t > 0$ . It is defined by two parameters: *reinforcement peak*  $0 < \mu \leq 1$  and *forgetting threshold*  $0 < \theta < \mu$ .

For clarity, we present here the basic CogSNet model; a more general version is defined in *Supplementary Information*, and Supplementary Eqs 6 and 7. This model processes an event happening at time  $t$  as follows. If the event involves a pair of unconnected nodes  $i, j$ , an edge  $(ij)$  linking these nodes is created. This edge is assigned weight  $w_{ij}(t)$  equal to the reinforcement peak value  $\mu$ . Otherwise, if the nodes involved are already connected, the forgetting function  $f$  is used to set the weight of the edge connecting these nodes to:

$$bw_{ij}(t) = \begin{cases} \mu, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ \mu + w_{ij}(t_{ij})f(t - t_{ij})(1 - \mu), & \text{otherwise,} \end{cases} \quad (1)$$

where  $t_{ij}$  is the time of the preceding event for this edge. Processing of each event ends with advancing  $t_{ij}$  to  $t$ . Initially,  $t_{ij}$  and  $w_{ij}(0)$  are set to 0. Eq. 1 and limit on value of  $\mu$  ensure that a weight of any edge is at most 1.

The weight  $w_{ij}(t)$  of an edge  $(ij)$  between two nodes at any user selected time  $t$  is computed as follows. Once all the relevant events up to time  $t$  are processed, we simply set  $w_{ij}(t) = w_{ij}(t_{ij}) * f(t - t_{ij})$ . If the result is less than the forgetting threshold  $\theta$ ,  $w_{ij}(t)$  is reduced to zero and the edge is removed. A threshold is needed with forgetting functions, such as power and exponential forgetting, that are positive for non-negative arguments. Otherwise, an edge would get positive weight at creation and would always stay positive, i.e., all created memory traces would never cease to exist. The reinforcement peak  $\mu$  defines the impact of an event on the weight of the edge relevant to this event. This peak is a global model parameter here. In principle, the peak can be adjusted according to the event or node type to allow for individualized event perception.

In general, the forgetting function  $f(\Delta t)$  over time interval  $\Delta t$  can be of any type (linear, power, logarithmic, etc.), but here, informed by work in the cognitive psychology of memory [10] we evaluate only two such functions: the exponential function  $f^{exp}$ , and the power function  $f^{pow}$  defined as:

$$f^{exp}(\Delta t) = e^{-\lambda \Delta t}. \quad (2)$$

$$f^{pow}(\Delta t) = \max(1, \Delta t)^{-\lambda}. \quad (3)$$

where  $\lambda$  denotes the forgetting intensity; typically  $\lambda \in [0, 1]$ . The use of max in the power function ensures that perception of events that happened less than a time unit ago is not changed by forgetting. The time unit in which the forgetting function is expressed scales the values of the parameters. Our experiments use one hour as the time unit.

To simplify optimal parameters search, we aggregate all three parameters into the *trace life time*  $L$  defining the time after which an unreinforced memory trace is forgotten, i.e., too hardly recalled. In the model,  $L$  is the time over which the forgetting function reduces the edge weight from  $\mu$  to  $\theta$  causing the edge to be removed, cf. Fig. 1. For the exponential forgetting function, Eq. 2, trace life time  $L^{exp}$  is:

$$L^{exp} = \frac{1}{\lambda} \ln \left( \frac{\mu}{\theta} \right), \quad (4)$$

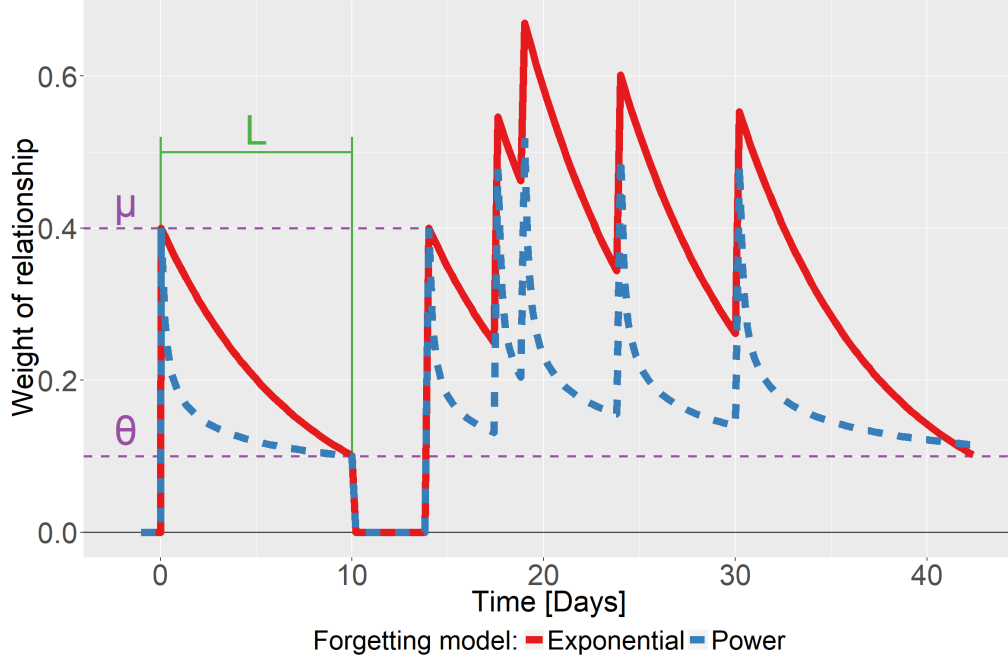


Figure 1: Dynamics of relations in CogSNet network with exponential and power functions and with parameters set to  $\mu = 0.4$ ,  $\theta = 0.1$ , and  $L=10$  days.

while for the power function, the formula is:

$$L^{pow} = \left(\frac{\mu}{\theta}\right)^{\frac{1}{\lambda}}. \quad (5)$$

In Fig. 2, we compare the CogSNet model with the previous proposals for representing temporal network dynamics. The most common approach for representing social network dynamics is to use interaction sequences [8]. Under this method, each event is time-stamped and the weights are added to the edges connecting nodes involved in this event, cf. Fig. 2(I) and 2(IIa). Moreover, a given edge is active (exists) only at a given time  $t$ . This is the most granular approach as it is capable of tracking all the events occurring between nodes while preserving the temporal order of events.

In contrast, a static binary network representation, as shown in Fig. 2(IIb), aggregates all events by making all edges time-independent. Consequently, an edge exists between a pair of nodes an event between these nodes occurred at least once in the whole observed period [7]. Such an edge representation throws away information on the temporal the ordering of events, making it impossible to study dynamic processes in static networks.

The incremental network solution accumulates events only up to the current time  $t$  of analysis. The classical approach, used early in [3,11,12], views a dynamic network as a series of time-ordered sequences of static graphs, see Fig. 2(IIc). More recently, this method was applied to modeling network and community evolution [14,18]. The drawback of this approach is that it does not preserve the ordering of interactions within time slices. Applying a simple frequency-based aggregation creates a frequency-based, FQ, metric, cf. Fig. 2(IId). Taking into account only a given number of the most recent events leads to the recency-based, RC, model, cf. Fig. 2(IIE). Both of these models are used here as baseline models.

Fig. 2(IIf) shows an example of dynamic social network generated using the CogSNet model. All other social network models presented in Fig. 2(IIa-e) can also be represented by CogSNet by setting appropriate parameter combinations to achieve, as needed, no decay, instant decay, and so forth. In this way CogSNet

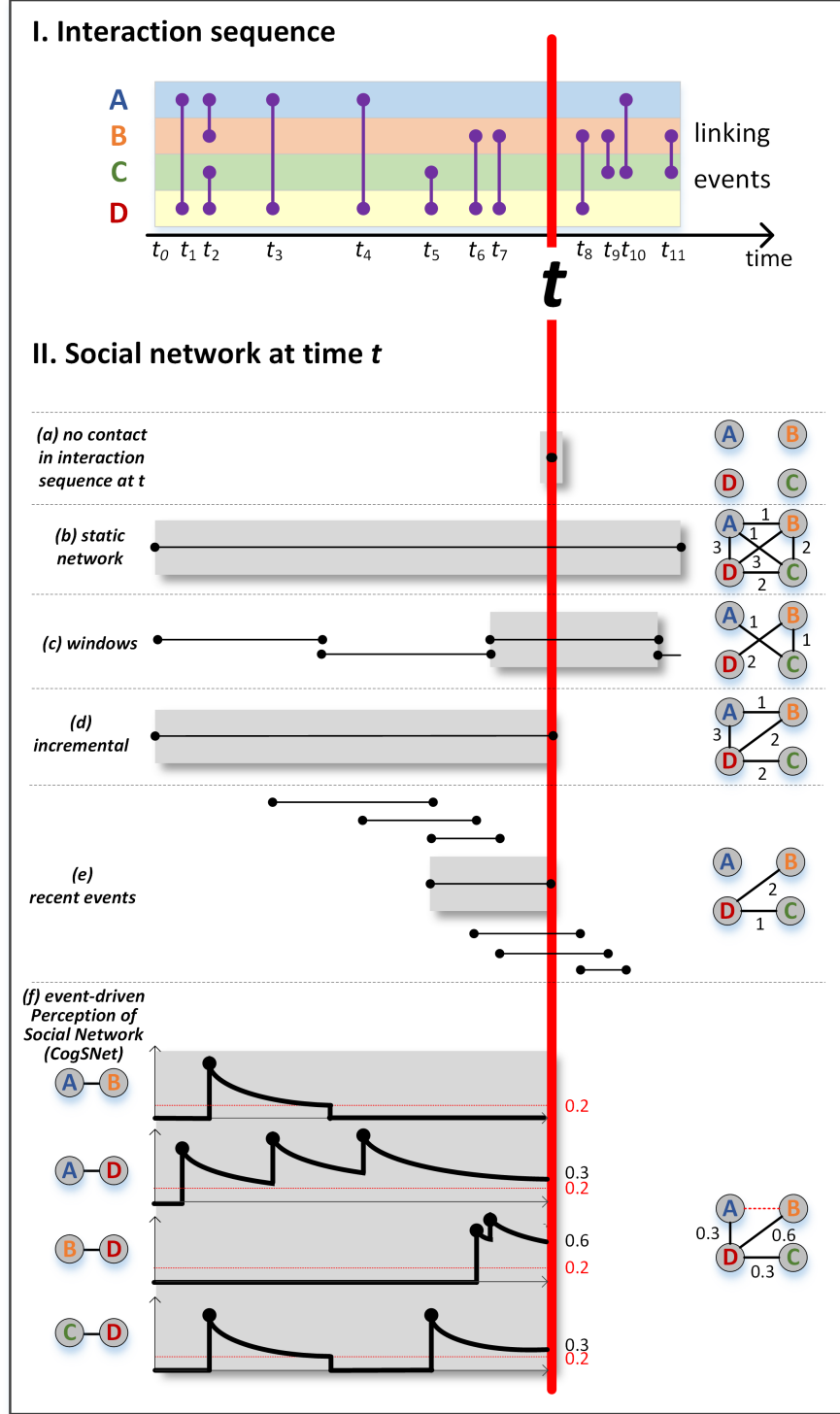


Figure 2: Various approaches to modeling dynamic networks and edge weighting (relation strengths) for the 4-node network at a given time  $t$ : **(I)** and **(IIa)** interaction sequence; **(IIb)** static (time-aggregated) network; **(IIc)** sliding windows; **(IId)** incremental network, all events from the beginning time  $t_0$  to the current time  $t$  are considered; the frequency of interaction in the period constitutes the frequency-based reference FQ; **(IIe)** network based on  $n = 3$  recent events, used for recency-based reference RC; **(IIf)** Cognition-based Social Network model, CogSNet, introduced here.

can be thought of as a universal generative dynamic model for temporal social networks, encompassing previous approaches as special cases.

## Results

In this section, we assess the empirical validity of the proposed model by estimating the accuracy with which the CogSNet model reproduces the dynamics of social relations in a dynamic social network. Our computational experiments use the NetSense dataset introduced in [16]. The data contain 6,290,772 human mobile phone communication events, including both calls and text messages. These are augmented by 578 surveys containing self-reports of top contacts.

We use this dataset to study the evolution of two coupled social networks of university students. The first is a behavioral network representing interactions between individuals in the form of the records of their mobile calls and text messages. The second one has perceptual edges defined by the personal nominations. These nominations are based on students’ perception of the corresponding relations as one of the top twenty most interacting peers in the surveys administered to participants. These surveys cover the first four semesters of the student’s college experience (beginning of Freshman year to the end of the Sophomore year). We compare the list of nominations predicted from the CogSNet network model purely from the communication event data with the list of nominations collected in a given survey. Fig. 3 shows an example of dynamic social network generated from a subset of NetSense data using the CogSNet model.

For comparison, we also implement three baseline models using the mobile phone communication data. The first is a *frequency-based*, FQ, model which orders the peers by the number of interactions, regardless of their time order. Under this model, the highest ranking peers whose number equals the number of the nominations listed in the corresponding survey are selected on the basis of their communication frequency. The *recency-based*, RC, model selects the top interacting peers who for each individual given a number of recent events. Finally, the *random sampling*, RND, model creates the list by randomly selecting peers from those who are recorded in the history of interactions of a given node.

To compare the performance of all the models, we use a *Jaccard* metric, Supplementary Eq. 8 in *Supplementary Information*, which measures the ratio of the number of nominations produced by the model that are also ground truth nominations listed in the corresponding survey divided by the number of unique nominations on both lists.

Fig. 4 shows the results of this comparison over the range of parameters corresponding to reported values for forgetting of one to 14 weeks. As reported in [5], the ability to recall information about social interactions starts to degrade after about one week. The experiments using NetSense dataset reveal that the performance is the highest when the forgetting of unreinforced memory traces happens after two weeks. The results remain satisfactory for forgetting thresholds as low as one week. With the two-week threshold, the model with either power or exponential forgetting is to be preferred, by the criterion of statistical significance, over any baseline model. Power forgetting is also significantly better than the exponential forgetting model. The performance of memory models using the power and exponential forgetting functions are similar to one another, but with a limited range of parameters, power forgetting tends to achieve higher Jaccard than the exponential function does [1]. We do not observe such superior performance of power forgetting here both functions have similar peak of Jaccard metric albeit for the different lifetime values. The detailed results are shown in *Supplementary Information*, see Supplementary Table 3.

When comparing the results of surveys with the states of the CogSNet network at the times of the surveys, the Jaccard is as high as 27.1%. The distant second is recency-based *RC* model which delivers much lower Jaccard of 17.8%.

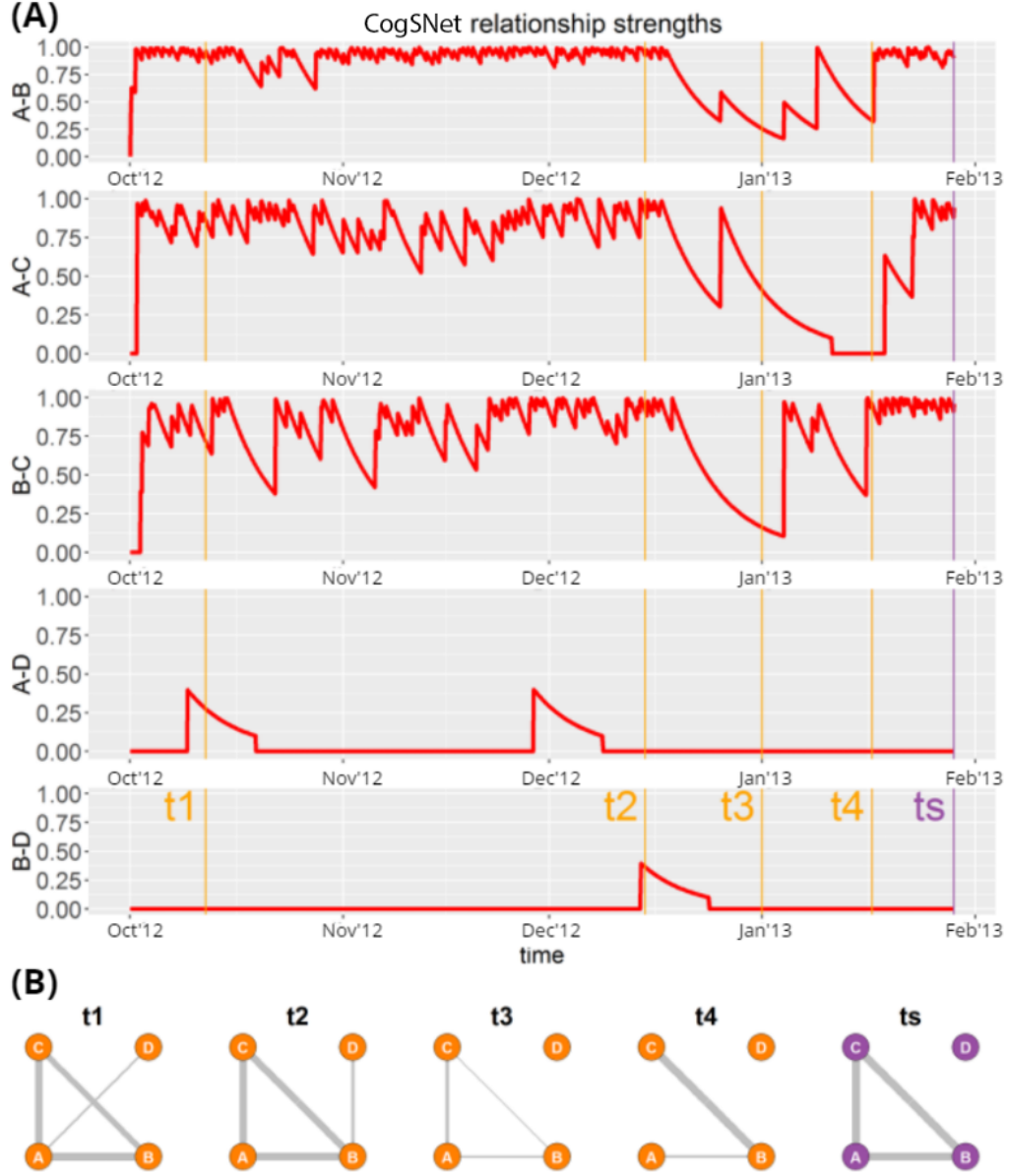


Figure 3: The four-node CogSNet network for the sample of real NetSense data,  $\mu = 0.4$ ,  $\theta = 0.1$ , and  $L=10$  days; nodes A, B, C, D correspond respectively to participants with ids 40997, 11360, 10841, and 1232. **(A)** Relation strengths according to the CogSNet model over 4-month period (one term). **(B)** Network snapshots at four time-stamps  $t_1$ - $t_4$  and at the survey time  $t_s$ .

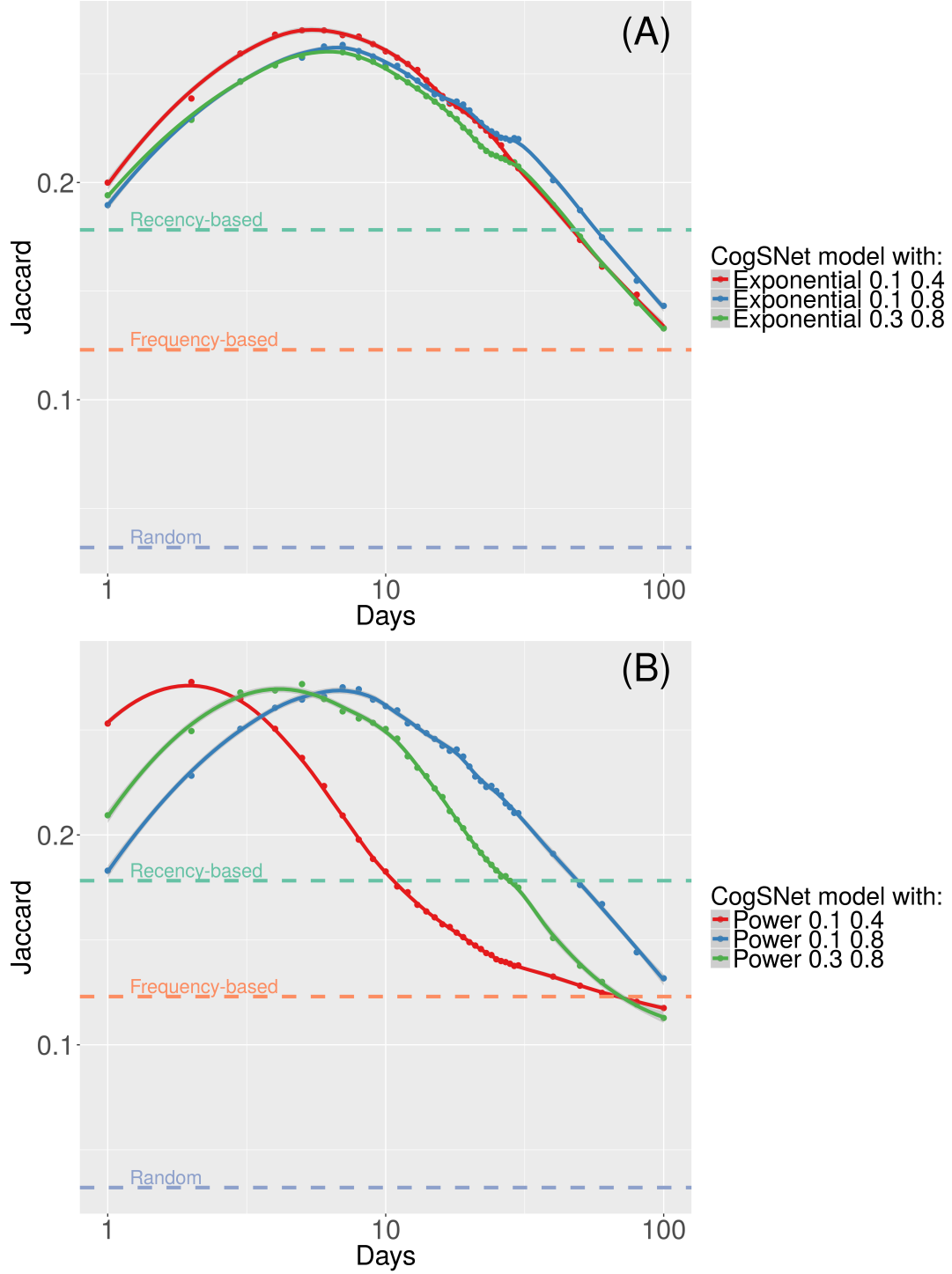


Figure 4: The plot with the Jaccard measure that presents the overlap between the set of peers in a network model with and set of peers in the survey. These plots are compared to the results achieved by the three baseline models: recency-based with the best results obtained with the number of recent events set to 400, frequency-based, and random. In the inset, the results are plotted just for the CogSNet running with various parameters for (A) exponential and (B) power functions.



## References

- [1] John R Anderson, Dan Bothell, Christian Lebiere, and Michael Matessa. An integrated theory of list memory. *Journal of Memory and Language*, 38(4):341–380, 1998.
- [2] Alan D Baddeley. The psychology of memory. *The essential handbook of memory disorders for clinicians*, pages 1–13, 2004.
- [3] Tanya Y Berger-Wolf and Jared Saia. A framework for analysis of dynamic social networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 523–528. ACM, 2006.
- [4] Rafael A Calvo and Sidney D’Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on affective computing*, 1(1):18–37, 2010.
- [5] Nathan Eagle, Alex Sandy Pentland, and David Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the national academy of sciences*, 106(36):15274–15278, 2009.
- [6] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. On graph problems in a semi-streaming model. In *International Colloquium on Automata, Languages, and Programming*, pages 531–543. Springer, 2004.
- [7] Petter Holme. Network dynamics of ongoing social relationships. *EPL (Europhysics Letters)*, 64(3):427, 2003.
- [8] Petter Holme and Jari Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.
- [9] Peter R Huttenlocher. *Neural plasticity: The effects of environment on the development of the cerebral cortex*. Harvard University Press, 2009.
- [10] Paul Jenkins, Giulia Earle-Richardson, David Tucker Slingerland, and John May. Time dependent memory decay. *American journal of industrial medicine*, 41(2):98–101, 2002.
- [11] Gueorgi Kossinets and Duncan J Watts. Empirical analysis of an evolving social network. *science*, 311(5757):88–90, 2006.
- [12] James Moody. The importance of relationship timing for diffusion. *Social Forces*, 81(1):25–56, 2002.
- [13] Donald A Norman. Memory and attention: An introduction to human information processing. 1969.
- [14] Vedran Sekara, Arkadiusz Stopczynski, and Sune Lehmann. Fundamental structures of dynamic social networks. *Proceedings of the national academy of sciences*, 113(36):9977–9982, 2016.
- [15] Sameet Sreenivasan, Kevin S Chan, Ananthram Swami, Gyorgy Korniss, and Boleslaw Karol Szymanski. Information cascades in feed-based networks of users with limited attention. *IEEE Transactions on Network Science and Engineering*, 4(2):120–128, 2017.
- [16] Aaron Striegel, Shu Liu, Lei Meng, Christian Poellabauer, David Hachen, and Omar Lizardo. Lessons learned from the netsense smartphone study. In *ACM SIGCOMM Computer Communication Review*, volume 43, pages 51–56. ACM, 2013.
- [17] John T Wixted and Ebbe B Ebbesen. On the form of forgetting. *Psychological science*, 2(6):409–415, 1991.

- [18] Jierui Xie, Mingming Chen, and Boleslaw K Szymanski. Labelrank: Incremental community detection in dynamic networks via label propagation. In *Proceedings of the Workshop on Dynamic Networks Management and Mining*, pages 25–32. ACM, 2013.

## Acknowledgment

Authors would like to thank Dr. Tomasz Kajdanowicz for his valuable remarks regarding statistical tests. This work was partially supported by the National Science Centre, Poland, by the following projects: 2015/17/D/ST6/04046 (RM) and 2016/21/B/ST6/01463 (PK, MK), by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 (the ARL Network Science CTA), by the Office of Naval Research (ONR) Grant No. N00014-15-1-2640, by the European Union’s Marie Skłodowska-Curie Program under grant agreement no. 691152, and by the Polish Ministry of Science and Higher Education under grant agreement no. 3628/H2020/2016/2. Calculations have been carried out using resources provided by Wrocław Centre for Networking and Supercomputing (<https://wcss.pl>), grant No. 177.

## Author contributions

All authors participated in the design of the research and computational experiments. RM and MK ran the computational experiments and collected the results. All authors participated in analysis of the results. All authors wrote and edited the manuscript.

## Competing financial interests

Authors declare no conflict of interest.

## Supplementary Information

### Formal definition of the CogSNet model

A social network can be represented by a graph  $G = (V, E)$ , where  $V = \{v_1, \dots, v_n\}$  denotes the set of  $n$  nodes and  $E = \{e_1, \dots, e_m\}$  is the set of directed  $m$  edges between pairs of nodes. Each edge  $e_{ij}$  from node  $v_i$  to node  $v_j$ ,  $i \neq j$ , is assigned weight  $w_{ij}(t)$  and represented by a triple  $(v_i, v_j, w_{ij}(t))$ .

The model evolves in discrete steps as follows. For each pair of nodes  $(v_i, v_j)$ , the system maintains two variables:  $t_{ij}$ , which represents the time of the most recent event for this pair of nodes, and  $c_{ij}$  which holds the count of events processed for this pair of nodes. Initially, both  $t_{ij}$  and  $c_{ij}$  are set to 0, as are the weights of all edges, i.e. for all pairs of nodes  $(v_i, v_j)$ ,  $w_{ij}(0) = 0$ .

When an event happens at time  $t$  in the modeled social network, it is processed in chronological order by the model. First, the weight of the corresponding edge is updated according to the following equation:

$$w_{ij}(t) = \begin{cases} \mu_{ijc_{ij}+1}, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ \mu_{ijc_{ij}+1} + w_{ij}(t_{ij})f(t - t_{ij})(1 - \mu_{ijc_{ij}+1}), & \text{otherwise,} \end{cases} \quad (6)$$

where:  $\mu_{ijk}$  is the value of reinforcement peak that results from the  $k^{th}$  event that impacts the edge  $(v_i, v_j)$ .

Here, the value of reinforcement peak  $\mu_{ijk}$  depends on the engagement and emotions invoked by the event that is either directly or indirectly related to the edge  $(v_i, v_j)$ . An example of an event indirectly related to this edge could be node  $v_i$  talking about node  $v_j$  or any situation that reminds node  $v_i$  about node  $v_j$ . The values of  $\mu$  can be individualized to node  $v_i$  perception of relation with node  $v_j$  at event  $k$ . The  $\mu$ 's values may also be dependent on event types:  $\mu_{ijk} \in \{\mu_1, \mu_2, \mu_3, \mu_4, \dots\}$ , e.g.,  $\mu_1 = 0.5$  for emails,  $\mu_2 = 0.55$  for phone calls,  $\mu_3 = 0.8$  for meetings,  $\mu_4 = 0.9$  for joint collaboration in projects, etc.

Finally, the processing of the current event updates both variables associated with the updated edge  $(v_i, v_j)$  as follows:  $t_{ij} = t, c_{ij} = c_{ij} + 1$ .

At any time  $t$  of the model evolution, the user can obtain the value of the weight of an arbitrary edge  $(v_i, v_j)$  by computing the following equation.

$$w_{ij}(t) = \begin{cases} 0, & \text{if } w_{ij}(t_{ij})f(t - t_{ij}) < \theta, \\ w_{ij}(t_{ij})f(t - t_{ij}), & \text{otherwise.} \end{cases} \quad (7)$$

### Experimental parameter space

To analyze the CogSNet model and compare its results to those obtained with the reference models, many combinations of the CogSNet parameters were tested. First, different values of  $\mu$ ,  $\theta$  and  $L$  were specified, then  $\lambda$  coefficient was calculated using Eqs 4 and 5 which were also used to define CogSNet parameters for experimental verification. The complete list of parameters used is listed in Supplementary Tables 1 and 2.

### Quality measures

The Jaccard for a single surveyed student participant  $v_i$  has been computed as follows:

$$Jaccard(v_i) = \frac{|V_i^{CogSNet} \cap V_i^{survey}|}{|V_i^{CogSNet} \cup V_i^{survey}|}, \quad (8)$$

where  $V_i^{survey}$  is the set of up to 20 peers enumerated in the survey by the participating student  $v_i$ ;  $V_i^{CogSNet}$  is the set of  $|V_i^{survey}|$  neighbors of this student in the CogSNet network with the largest non-zero weights on edges to this student on the day on which the given survey was administered.

Table 1: Setting of parameters for power forgetting function in computational experiments

$\lambda$ dependence on $\mu, \theta$ and $L$ for power forgetting			
Life time $L$ days	$\mu = 0.4; \theta = 0.1$	$\mu = 0.8; \theta = 0.3$	$\mu = 0.8; \theta = 0.1$
1	0.43621	0.30863	0.65431
2	0.3581	0.25337	0.53716
3	0.32415	0.22934	0.48623
4	0.30372	0.21489	0.45558
5	0.28957	0.20487	0.43435
6	0.27894	0.19736	0.41841
7	0.27055	0.19142	0.40583
8	0.26368	0.18656	0.39552
9	0.2579	0.18247	0.38685
10	0.25294	0.17896	0.37942
11	0.24862	0.1759	0.37293
12	0.2448	0.1732	0.3672
13	0.24139	0.17079	0.36208
14	0.23831	0.16861	0.35747
15	0.23552	0.16663	0.35328
16	0.23297	0.16483	0.34945
17	0.23062	0.16317	0.34592
18	0.22844	0.16163	0.34267
19	0.22643	0.1602	0.33964
20	0.22455	0.15887	0.33682
21	0.22278	0.15762	0.33418
22	0.22113	0.15645	0.3317
23	0.21957	0.15535	0.32936
24	0.2181	0.15431	0.32716
25	0.21671	0.15333	0.32507
26	0.21539	0.15239	0.32309
27	0.21414	0.15151	0.3212
28	0.21294	0.15066	0.31941
29	0.2118	0.14985	0.3177
30	0.21071	0.14908	0.31606
40	0.20188	0.14283	0.30282
50	0.19553	0.13834	0.29329
60	0.19062	0.13487	0.28594
80	0.18337	0.12974	0.27506
100	0.17811	0.12602	0.26717

Table 2: Setting of parameters for exponential forgetting function in computational experiments

$\lambda$ dependence on $\mu, \theta$ and $L$ for exponential forgetting			
Life time $L$ (days)	$\mu = 0.4; \theta = 0.1$	$\mu = 0.8; \theta = 0.3$	$\mu = 0.8; \theta = 0.1$
1	0.05776	0.04087	0.08664
2	0.02888	0.02043	0.04332
3	0.01925	0.01362	0.02888
4	0.01444	0.01022	0.02166
5	0.01155	0.00817	0.01733
6	0.00963	0.00681	0.01444
7	0.00825	0.00584	0.01238
8	0.00722	0.00511	0.01083
9	0.00642	0.00454	0.00963
10	0.00578	0.00409	0.00866
11	0.00525	0.00372	0.00788
12	0.00481	0.00341	0.00722
13	0.00444	0.00314	0.00666
14	0.00413	0.00292	0.00619
15	0.00385	0.00272	0.00578
16	0.00361	0.00255	0.00542
17	0.0034	0.0024	0.0051
18	0.00321	0.00227	0.00481
19	0.00304	0.00215	0.00456
20	0.00289	0.00204	0.00433
21	0.00275	0.00195	0.00413
22	0.00263	0.00186	0.00394
23	0.00251	0.00178	0.00377
24	0.00241	0.0017	0.00361
25	0.00231	0.00163	0.00347
26	0.00222	0.00157	0.00333
27	0.00214	0.00151	0.00321
28	0.00206	0.00146	0.00309
29	0.00199	0.00141	0.00299
30	0.00193	0.00136	0.00289
40	0.00144	0.00102	0.00217
50	0.00116	0.00082	0.00173
60	0.00096	0.00068	0.00144
80	0.00072	0.00051	0.00108
100	0.00058	0.00041	0.00087

## Recency-based baseline

The recency-based baseline model was tuned by finding the best parameter representing the number of the most recent events to be taken into account in computing the performance. The resulting value of 400 events was used in Fig. 4 for computing the results for the recency-based baseline model.

	Power	Exponential	Recency	Frequency	Random
Power	—	***	***	***	***
Exponential	0.0000	—	***	***	***
Recency	0.0000	0.0000	—	***	***
Frequency	0.0000	0.0000	0.0000	—	***
Random	0.0000	0.0000	0.0000	0.0000	—

Table 3: Numerical and symbolic p-values adjusted by Simes-Hochberg step-up method of non-parametric pairwise post-hoc Nemenyi test comparing Jaccard of methods for surveys taken over semesters 1-4. The results show how each method in the leftmost column performs versus the method in the first row. The part of table under the diagonal shows numerical values of the test, while the part above the diagonal presents p-values coded as: \*\*\* for  $p < 0.00005$  which represents high confidence in the performance differences between the methods since typically values of  $p > 0.05$  are considered indicative of non-significance.

## Statistical tests

To confirm that there is a statistical difference between the proposed CogSNet model and baselines, we performed a number of statistical tests. The results are presented in Table 3 which shows that CogSNet approach is statistically significantly better from all baseline methods. Moreover, power forgetting performs statistically significantly better than exponential forgetting.

## NetSense dataset

The dataset consists of two parts. The first includes the time-stamps and duration/length of phone calls and text messages collected for each student participating in the study. Each student phone device recorded all connections/messages, including those to the phones of people outside of the test group, so recording was done on both sides of communication, by sending and receiving calls/messages, if both belonged to the study.

The second part of the dataset includes surveys containing peers enumerated by the participants at the end of each term in response to the following question: *"In the spaces below, please list up to 20 people (friends, family members, acquaintances, or other people) with whom you spend time communicating or interacting"*.

Due to the fact that the dataset covers millions of mobile phone interactions, some issues discussed below arose during the preliminary data analysis phase. When the students phoned each other and the call was not answered, the recipient's voice mail was reached, if it was enabled. Such case has been recorded as a regular phone call even if it has not been actually answered by the recipient.

Since events involving each student possessing the mobile phone were recorded independently from other mobile phones, some inconsistencies arose among which the most important was that not in all cases the event was recorded on both sides. Since usually a couple of seconds pass between sending text message and receiving it by the recipients, we identified the same message by its length. However, sometimes the

recorded lengths were different, so in such cases we were forced to make somewhat arbitrary decision based on the length difference whether the message was the same or separate one for which the corresponding event on the other side of communication was not recorded.

Lastly, for some surveys, the students were asked to enumerate in any order up to twenty peers with whom they interacted recently. However the system recorded each name with the accompanying number representing name's position on the list. For some surveys there were some holes in this numbering, so either the student did not list the peers accordingly (left some blank lines in between names) or partial results were not recorded. This is why we consider listed peers as the sets without any implied ordering.

The NetSense dataset used for the model validation is available for other researchers; the readers interested in accessing the data should contact the co-author, Professor Omar Lizardo.